

AI Is Tricky

TEXAS Grand Challenges



[The Grand Challenge of Ethics and AI: A Fireside Chat with Tim Hwang, October 7, 2019](#)

When people hear the term artificial intelligence or AI, there may be a tendency to think of the latest “Terminator” movie or fall for the hype of menacing machines controlling our existence and destroying humanity.

[Tim Hwang](#), a lawyer, writer, and researcher working at the intersection of emerging technologies and society, says this isn’t realistic; machine learning and AI aren’t as they’re portrayed in the movies. Hwang is the former director of the [Harvard-MIT Ethics and Governance AI Initiative](#) and has previously served Google’s global public policy lead on AI. In a recent [IEEE Spectrum](#) article, he says, “I think that a lot of the people who work on machine learning on a day-to-day basis are pretty humble about the technology because they’re largely confronted with how frequently it just breaks and doesn’t work.”

[Hwang researches and writes about the social implications of artificial intelligence.](#) “After seeing how machine learning was being done within the tech industry, it was obvious that there is a great need to bolster voices outside these corporations working to ensure that these technologies are accountable to the public,” he explains.

At the same time, breakthroughs in AI have led some groups to study the ethics and governance of the technology. “People tend to forget, but AI was considered a dead end for a long time. Machine learning, in particular, had failed to live up to some of the grand promises that were made about it in the 1960s and 1970s,” says Hwang, “but the pendulum has swung as changes in the accessibility of data and computing power have shown that these methods really work.” Now Hwang and others working in AI are asking, what are the risks, and are there contexts in which we shouldn’t use them?

AI is not the Terminator. “It’s not going to climb out of your computer and destroy you.” Negligence and incompetence around the use of AI are the real threats.

He emphasizes, though, that those questions need to go even deeper now. We should be asking ourselves: how do we manage public and private power of AI? How do we continue to strive to see what’s behind the curtain when small groups hold so much of the research? How is predictive policing designed and deployed so AI is a benefit, not a detriment to society? He also stresses that we need to determine what our hard tradeoffs are and create legitimate processes to help manage them.

“AI is tricky,” Hwang concedes. “The algorithms used for finding kittens in an image are the same as those for targeting drones.” He says we need to decide which domains have the greatest need and focus on those, such as healthcare and surveillance. Since many of the leading academic and industrial labs working on machine learning are concentrated in China or the United States, international norms are another critical domain. How do all countries benefit equally from AI? That may require creating a technical space in government that adheres to a broad ethical initiative.

When asked about deepfake AI technology and its relationship to today’s disinformation epidemic, Hwang says he’s a skeptic. “I don’t think they’ll have as big an impact as many people believe and may be a distraction” to the real issues facing AI technology. “We need to keep thinking about how the technical space compliments and conflicts with broader social forces and how we’re going to keep those aligned.”

He adds, jokingly, that AI is not the Terminator. “It’s not going to climb out of your computer and destroy you.” Negligence and incompetence around the use of AI are the real threats.

[Tim Hwang will share his perspectives about ethics and AI on October 7, 2019, on the UT Austin campus.](#) This event is open to the public and will help launch Good Systems as UT’s third grand challenge: an 8-year research initiative with the goal of designing AI technologies that benefit society.

Please join us on this journey.

[Good Systems](#) is a research grand challenge at The University of Texas at Austin. We’re a team of information and computer scientists, robotics experts, engineers, humanists and philosophers, policy and communication scholars, architects, and designers. Our goal over the next eight years is to design AI technologies that benefit society. Follow us on [Twitter](#), join us at our [events](#), and come back to our [blog](#) for updates.